V. Databases
B.iv. Science Databases

Robert Schaefer – GSSC Software Lead

Outline

- Requirements and Documents
- Data Products
- Databases
- Summary

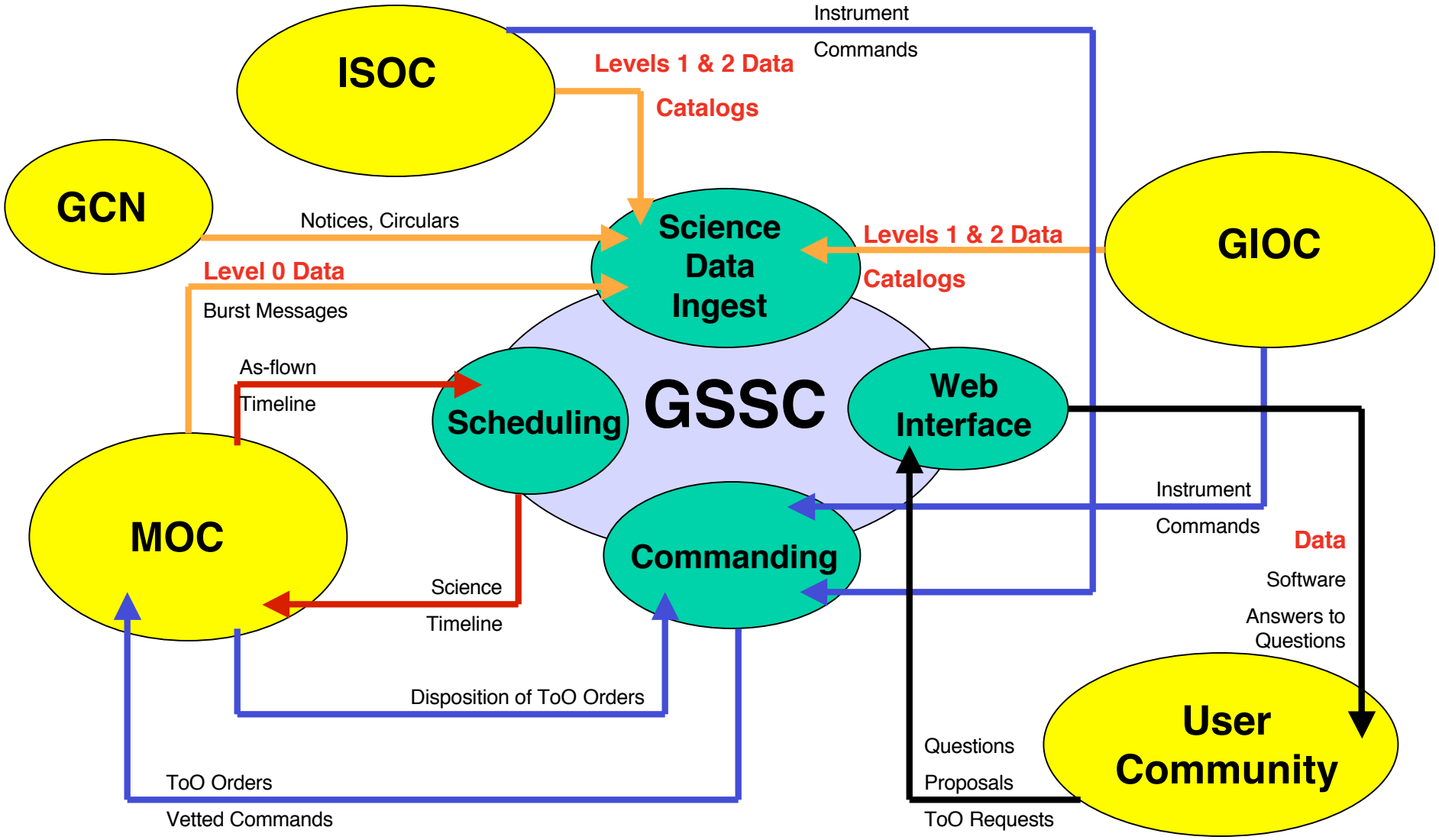
Documents

- Applicable Documents
  - GSSC Design Document GSSC-0003
  - Science Data Products ICD
  - Operations Data Products ICD
  - GSSC-HEASARC MOU
- Other applicable documents include
  - Project Data Management Plan (433-PLAN-0009)

Science Database Requirements

- Maintain databases for all the data products we receive from the MOC or IOCs. (FRD 5.7.2)
- These databases will be physically connected to the HEASARC computer system (FRD 5.7.3)
- Most databases will be accessed though via the web in accordance with the Mission data policies (FRD 5.7.4). Data served includes:
  - Processed data from the IOC's and GSSC
  - Catalogs
  - Calibration data
- The GSSC data and software will be transferred to the HEASARC by the end of the mission. (FRD 5.7.5, 5.7.6)
- RFA 12 - Two separate data interfaces GSSC and HEASARC - all GLAST data will be accessible through HEASARC's Browse. No need for someone to learn a separate interface for a standard analysis.

Data Flows and the GSSC Interaction



Science Data Products

Science Data Products - Browse			
Data Product	**Origin**	**Database**	**ICD**
GBM Daily Data Products (CTIME, CSPEC, housekeeping, and attitude files)	GIOC	Browse	GS-001, GS-002, GS-004, GS-006
GBM Burst data products (CTIME, CSPEC, TTE, TRIGDAT, DRMs, and background files)	GIOC	Browse	GS-101, GS-102, GS-104, GS-107, GS-108
GBM Burst Catalog	GIOC	Browse	GS-206
GBM Burst Spectra Catalog	GIOC	Browse	GS-306
GBM Trigger Catalog	GIOC	Browse	GS-207
LAT Burst Catalog	ISOC	Browse	LS-009
LAT Transient Data	ISOC	Browse	LS-007
LAT Point Source Catalog	ISOC	Browse	LS-008
LAT Interstellar Emission Model	ISOC	Browse	LS-010
Pulsar Ephemerides	GSSC	Browse	SS-002

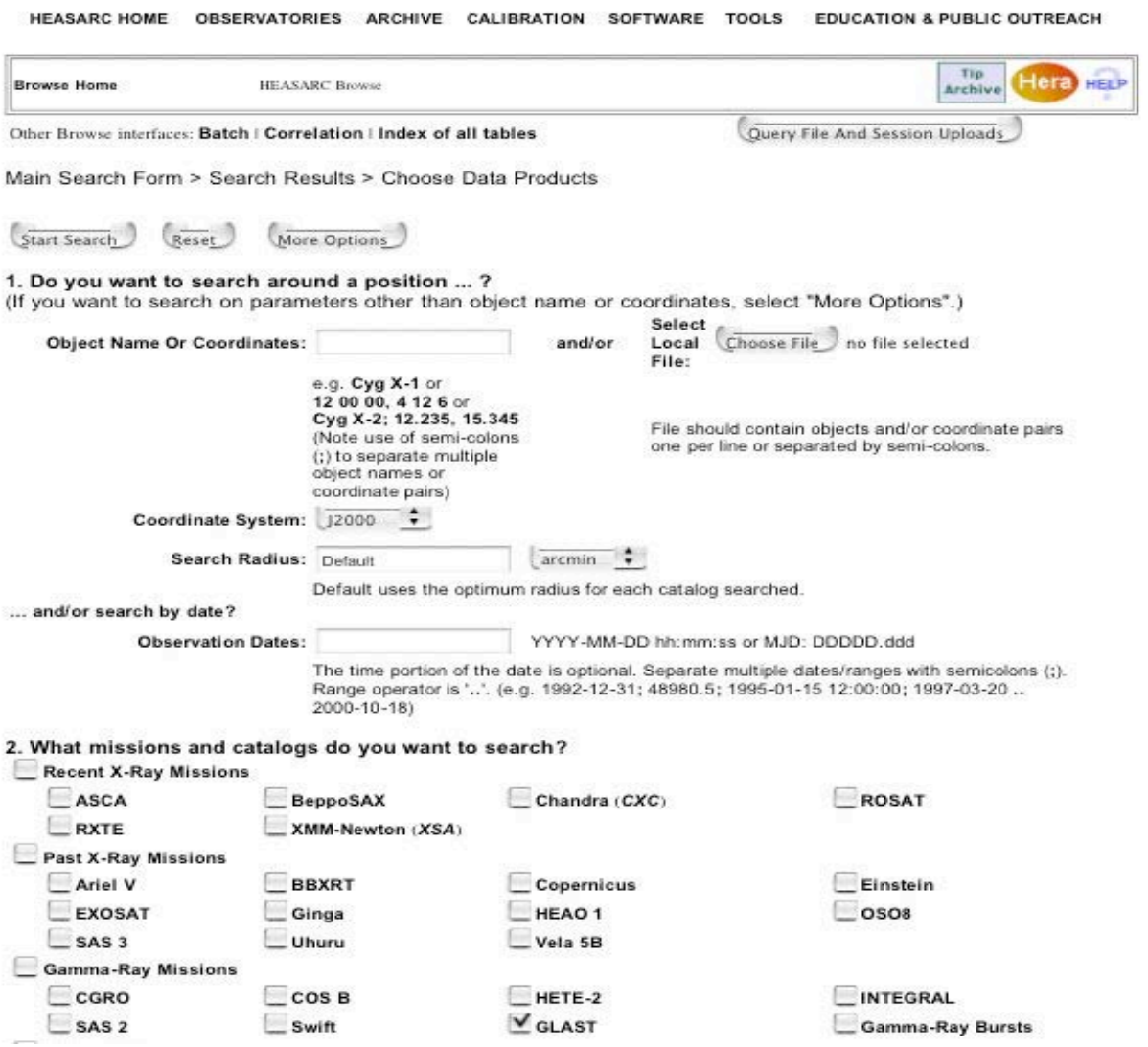
Science Databases

Science Data Related Products			
Data Product	**Origin**	**Database**	**ICD**
GBM Calibration Files	GIOC	CALDB	GS-005, GS-007
LAT IRF	ISOC	CALDB	LS-004
LAT Events data	ISOC	D1 database	LS-002
LAT Pointing and Livetime History Data	ISOC	D2 database	LS-005
Level 0 Data	MOC	Level 0 Database (D0)	Ops
Data Re-transmission Requests	GSSC	Internal Database	Ops
Data Processing	GSSC	Internal Database	N/A
Process Tracking	GSSC	Internal Database	N/A
Data Products	GSSC	Internal Database	N/A
Issues (Issue Tracker)	GSSC	Internal Database	N/A

Science Databases

- Browse - HEASARC Multi-Mission Archive
  - All relevant science data from HEA missions (data, catalogs, timelines, and proposals)
  - Well defined and documented system for ingest, data indexing
- CALDB - HEASARC
  - Flexible structure for instrument teams to maintain calibration data relevant for scientific analyses
  - Interface for tools to get files relevant for data sample
- Level 0
  - Internal Archive consists of indexing database (MySQL) and file archive
- MySQL Databases
  - Data Tracking, Processing, Issue tracking.
  - Databases will track a variety of information including delivery dates/times, version numbers, reasons for new versions, etc.

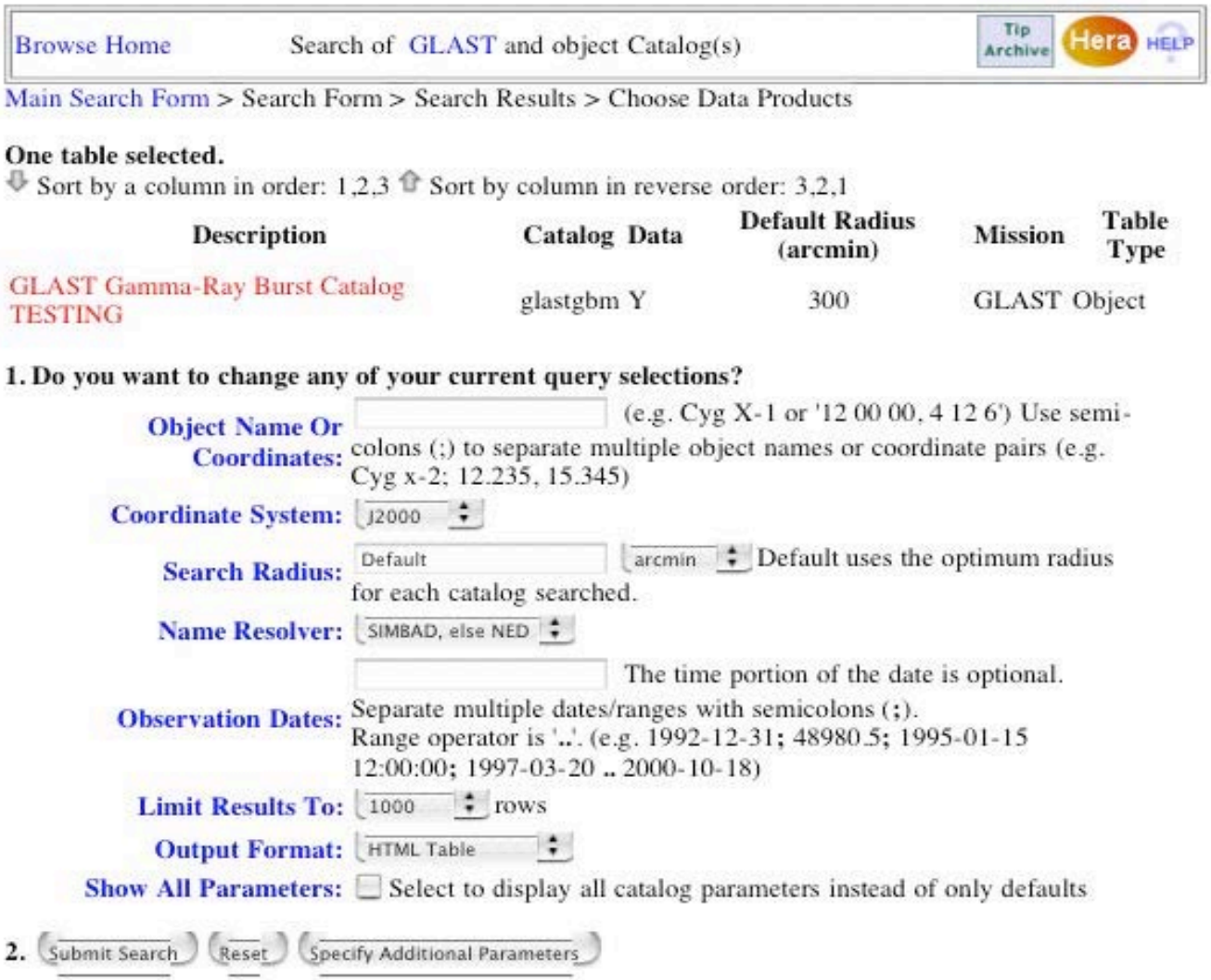
Browse Interface is Multi-Mission



Getting Data into Browse

- Data Ingest pipeline
  - Receives data, extracts metadata, creates Browse “tdat” file from metadata sent via DTS to HEASARC.
  - Copies data products to HEASARC archive area
  - Ingested catalogs will be linked to data products -
    - LAT catalogs will link to our event search engine
  - HEASARC database system will update tables on a shorter (~hours) with new information
  - Test tables have already been ingested into HEASARC Browse.
- Will piggyback on HEASARC data backup system.

GLAST Test Data in Browse



Browse Home Search of GLAST and object Catalog(s) Tip Archive Hera HELP

Main Search Form > Search Form > Search Results > Choose Data Products

One table selected.

Sort by a column in order: 1,2,3 Sort by column in reverse order: 3,2,1

Description	Catalog Data	Default Radius (arcmin)	Mission	Table Type
GLAST Gamma-Ray Burst Catalog TESTING	glastgbm Y	300	GLAST	Object

1. Do you want to change any of your current query selections?

Object Name Or Coordinates: (e.g. Cyg X-1 or '12 00 00, 4 12 6') Use semi-colons (;) to separate multiple object names or coordinate pairs (e.g. Cyg x-2; 12.235, 15.345)

Coordinate System: J2000

Search Radius: Default arcmin Default uses the optimum radius for each catalog searched.

Name Resolver: SIMBAD, else NED

Observation Dates: The time portion of the date is optional. Separate multiple dates/ranges with semicolons (;). Range operator is '..'. (e.g. 1992-12-31; 48980.5; 1995-01-15 12:00:00; 1997-03-20 .. 2000-10-18)

Limit Results To: 1000 rows

Output Format: HTML Table

Show All Parameters: Select to display all catalog parameters instead of only defaults

2. Submit Search Reset Specify Additional Parameters

CALDB

- CALDB
  - Instrument Teams responsible for determining data structure for CALDB information (in OGIP compliant FITS files).
  - New files delivered to GSSC will trigger testing procedure.
    - GSSC scientists will verify that files are valid.
    - Files copied to CALDB directory tree when approved.
  - Will be backed up with HEASARC backup system

Level 0

- GSSC responsible for permanently archiving Level 0 data.
- Normally, GSSC turns over the data to the HEASARC by mission end. Data that are little used end up at the NSSDC.
- However, most likely neither the GSSC nor the HEASARC will ever access these data; the HEASARC does not want it.
- RFA #7 asked why the IOCs, the GSSC and the NSSDC all archive the Level 0 data. What the IOCs do is an issue for the instrument teams; since they process the data, it is understandable that they keep Level 0 data around.
- Therefore, the GSSC will store Level 0 data for ~ one year before giving it to the NSSDC for deep storage. At the NSSDC the data will NOT be online for immediate electronic access.

MySQL

- Powerful, high performance Open Source Relational Database Management System
- Provides a suite of tools including admin and backup tools. (GSSC will backup MySQL databases)
- APIs in many languages (perl, C, etc.)
- In use in HEASARC

Summary

- Data Products available through Browse - stored in HEASARC compatible format
- No need to transfer most products to HEASARC - as they will already be in HEASARC system. (D1,D2 will be turned over later)
- Maintain year-long level 0 archive
- Using MySQL - open source standard database (already used by the HEASARC) for internal databases.